
E. Glossary

A

address unit

This collection unit is a person or group of persons living at the same address at the time of the interview. The address unit may consist of one person living by himself or herself, a group of unrelated individuals, or one or more families.

allocation flag

See imputation flag.

B

C

CAI (computer-assisted interviewing)

A method of interviewing in which a computer is used as the data collection instrument.

CAPI (computer-assisted personal interviewing)

A method of interviewing in which field representatives use a laptop computer to collect data during in-person interviews. In SIPP, the field representatives also periodically use the laptop computers during telephone interviews conducted from their homes.

cold-deck matrix

The matrix of starting values that constitutes the first step in the hot-deck imputation procedure. The matrix values can be determined a priori from information external to the current file being processed or can be determined from reported information from the current file.

control card

In the paper instrument for SIPP, a mechanism for carrying demographic and case management information forward from one wave to the next for each sample member.

core content

Questions asked at every SIPP interview. They cover demographic characteristics, work experience, earnings, program participation, transfer income, and asset income.

core wave files

Files containing the core data from one wave of interviews.

cross-sectional

Pertaining to data collected for a single time period from a representative sample. In SIPP hot-deck imputation procedures, cross-sectional refers to current-wave data.

Current Population Survey (CPS)

A labor force survey sponsored jointly by the Census Bureau and the Bureau of Labor Statistics that is used to compute the government's official monthly unemployment statistics along with other estimates of labor force characteristics.

D

data dictionary

Contains information about the file structure and the names, locations, and contents of all variables in a microdata file.

data editing

The use of related information to replace missing or inconsistent data in the survey.

departure noninterview

This type of noninterview occurs when someone was a member of a SIPP interviewed household during the 4-month reference period but was no longer a household member on the date of the interview.

E

F

family

Two or more people who are living together and are related by blood, marriage, or adoption.

FERRET

An on-line data access tool available on the SIPP Web site. SIPP data are available on FERRET beginning with the 1992 longitudinal panel.

following rules

SIPP rules that guide which original sample members continue to be interviewed should they move.

full panel files

Files containing all data for every person who was a member of a SIPP panel at any time during the life of that panel.

G

general income

Any type of income except earnings and asset income.

geographic (GRIN) codes

Codes that identify where each sample household is located and permit linkage to a file that contains a full set of geographic codes for different kinds of areas. This level of geography is not available on the public use files.

group quarters

Noninstitutional living quarters, such as rooming and boarding houses, college dormitories, convents, and monasteries. These do not constitute households and are often treated differently from households.

H

hot-deck matrix

The matrix used in all but the first stage of hot-deck imputation. As cold-deck values are replaced with information from the current wave, the resulting array of cells constitutes the hot-deck matrix.

hot-deck procedure

The statistical method used to impute items missing from the core questionnaire and topical modules. This procedure replaces missing item data in a wave with nonmissing values from similar interviewed cases. The imputation method can be a purely cross-sectional procedure of locating donors from the current file on the basis of characteristics reported in this wave, or it can be a longitudinal procedure of locating donors from the prior wave on the basis of characteristics reported at that earlier time for items missing in the current wave.

household

People living in a housing unit at the time of the interview. SIPP infers households from the interviews conducted at each address.

household-level noninterviews

See *household nonresponse*.

household nonresponse

Nonresponse that occurs when the interviewer either cannot locate a household or cannot interview any of its adult members. See *Type A*, *Type B*, *Type C*, and *Type D noninterviews*.

household reference person

See *reference person*.

housing unit

Living quarters with its own entrance and cooking facilities.

I**imputation**

The most common method for handling missing data in SIPP. Imputation replaces missing values with statistical estimates that are based on the best relevant information available.

imputation flag

An imputation flag is associated with each core questionnaire item subject to statistical imputation and indicates whether information has been imputed.

in-sample variables

See *monthly interview status variables*.

in scope

Being part of the survey universe.

interview month

The month during which the interview takes place.

item nonresponse

A source of missing data that occurs when a respondent does not answer one or more questions, even though most of the questionnaire is completed.

J**K****L****logical imputation**

See *data editing*.

longitudinal

Pertaining to data collected at different times over an extended period from a representative sample. In SIPP hot-deck imputation procedures, longitudinal refers to previous-wave data.

M

merged households

Households created either when two separate sampling units, each containing original sample members, are merged together, perhaps because of a marriage, or when a household splits into two new households and later the households recombine.

microdata files

Data files containing information at the person, family, or household level. For SIPP, they include the core wave files, topical module files, and full panel files.

missing item data

Data that are missing for one or more individual questions or variables, but the observation has sufficient reported information to be classified as interviewed.

missing waves

Waves in which a respondent has no data, although data are present for other waves.

monthly interview status variables

Variables that indicate whether a person was in sample in a particular month, and whether a person was in sample in the interview month. They are known as the PP-MIS variables.

mover

An original sample person who moves during the life of the panel.

N

National Longitudinal Survey (NLS)

Collects data on current labor force and employment status, work history, and characteristics of the current or last job.

non-self-representing (NSR) primary sampling units (PSUs)

Smaller PSUs that must be grouped with similar PSUs from the same region in order to form strata for sampling. This level of geography is not available on the public use files.

O

original sample members

All people who were interviewed in the first wave of the panel and any children subsequently born to or adopted by them.

oversampling

Sampling that involves selecting certain groups or units with higher probabilities than others, resulting in the oversampled group having greater representation than occurs in the population from which it was drawn.

P

P-70 reports

Primary source for published estimates from the SIPP. These reports can be obtained from the SIPP Web site or from the Census Bureau.

panel

Refers both to a new sample that is introduced periodically in the SIPP and to the full collection of information for that sample. For example, the 1996 Panel refers to both the sample introduced in 1996 and the 12 waves of interviews conducted with that sample.

panel nonrespondents

Persons for whom an interview is missing for a wave.

Panel Study of Income Dynamics (PSID)

A nationally representative, longitudinal survey of the U.S. population, conducted by the University of Michigan. The focus of the survey is economics and demographics, especially income sources and amounts, employment, family composition changes, and residential location.

Partial panel files

Longitudinal files to be released by the Census Bureau prior to the conclusion of the 1996 Panel because of the 4-year duration of the 1996 Panel.

person-level noninterviews

This type of noninterview occurs when data are collected for at least one member of a household, but are missing for one or more other sample persons within that household.

person-month files

Microdata files containing a record for each person in a wave, for each month of the reference period the person was in the sample.

person nonresponse

Nonresponse that occurs when at least one person in the household is interviewed, while at least one other person is not. See *Type Z noninterview*.

primary family

Family containing the household reference person and related individuals.

primary individual

A household reference person who lives alone or lives with only nonrelatives.

primary sample members

See *original sample members*.

primary sampling units (PSUs)

Geographic units based on Census data and used in developing the SIPP sample. This level of geography is not available on the public use files.

program units

The group of individuals which constitutes one case, as defined by a particular benefit program. In SIPP, program units apply to health insurance and transfer programs and are identified for programs in which a case can consist of more than one person.

proxy interviews

Interviews taken on behalf of a sample member who is unable to answer.

public use microdata files

Data files that have been prepared by the Census Bureau for public use. These files have already been processed to impute missing data, to edit data for confidentiality, and to provide weights. Microdata files are available from the Census Bureau or on-line from the SIPP Web site.

Q**R****random carryover method**

Longitudinal imputation procedure used to impute missing wave data.

1996 Redesign

A revamping of SIPP in order to improve the quality of estimates and to make the data more useful to analysts.

reference months

The months that constitute the reference period for a wave. The months vary for different rotation groups.

reference period

The 4 calendar months preceding the month of interview. The reference period is a different calendar period for each rotation group.

reference person

An owner or renter of record who can reasonably be expected to answer questions about the household in general and about other household members should they be unavailable for interview. All people in the household are listed according to their relationship to the reference person.

related subfamily

A married couple and dependents or parent-child family related to the reference person but not including him or her. An example would be the reference person's daughter and son-in-law.

rotation group

A subsample containing roughly one-quarter of the sample members. One rotation group is interviewed each month of a 4-month wave.

S

sample attrition

Loss of sample members. Sample attrition rates decline over time, but total attrition numbers increase.

seam effect

The tendency of respondents to report a disproportionate number of changes as occurring at the "seam" between the fourth month of one wave and the first month of the following wave.

secondary families

Two or more people living in the same household who are related to each other but not to the household reference person.

secondary individual

An individual who is neither a household reference person nor a relative of any other people in the household.

secondary sample members

People living with original sample members.

self-representing (SR) primary sampling units (PSUs)

Larger PSUs that do *not* have to be combined with other PSUs in order to form strata for sampling. This level of geography is not available on the public use files.

sequential hot-deck procedure

See *hot-deck procedure*.

short waves

Waves that contain three rotation groups instead of the standard four.

skip patterns

Mechanisms embedded in the survey that allow the interviewer to skip over irrelevant questions and call up the next relevant question.

source and accuracy statement

A statement included with the technical documentation that accompanies public use files; it contains detailed information about weights on the files, when and how to make adjustments to the weights, and how to use generalized variance procedures to compute standard errors for some common types of estimates. It also includes cautions for users about sources of nonsampling error.

Survey of Program Dynamics (SPD)

An offshoot of SIPP that began recontacting members of the 1992 and 1993 Panels, with data collection to continue through 2001 in order to collect 10 years of data.

Surveys-on-Call

An on-line data access tool available on the SIPP Web site. Surveys-on-Call allows users to define microdata extracts from SIPP public use files through the 1993 Panel.

T

technical documentation

Information that accompanies microdata files and that includes a description of file contents, a glossary, codes, a data dictionary, a source and accuracy statement, and a copy of the core questions for the panel in question.

time-in-sample effect

Tendency of sample members to “learn” the survey over time, possibly resulting in altered responses.

topcoding

Practice of recoding income variables to protect against the possibility that a user might recognize the identity of a SIPP respondent with very high income. Incomes exceeding a maximum value are recoded to that maximum value or to a mean of responses in excess of that value.

topical content

Questions that are not repeated in every wave. They cover a wide range of topics and can occur once or more than once in a panel. The questions are grouped into modules by topic.

topical module files

Files containing all topical module data from the wave in question.

topical modules

Collections of questions asked periodically, but not at every interview, about various topics that might be outside the range of the core content.

topical module imputation procedure

Missing data in topical modules are imputed using the same hot-deck procedure used to impute missing data in the core questionnaire.

Type A noninterview

Households that are occupied by people eligible for interview but for which no interview is obtained.

Type B noninterview

A household noninterview that occurs when the address unit is vacant or in some way unfit for residence.

Type C noninterview

In Wave 1, a household noninterview that occurs when the housing unit has been demolished or converted to some other use; in subsequent waves, a household noninterview that occurs when all sample members in a household are outside the scope of the survey, for example, deceased, living abroad, living in institutions, or living in armed forces barracks.

Type D noninterview

Households or people who have moved to an unknown address, or who have moved more than 100 miles from the nearest field representative and for whom no telephone interview is conducted. This type of noninterview applies only to Wave 2 and beyond.

Type Z imputation

Procedures used to impute missing data for Type Z noninterviews and for situations when a person was in sample early in the wave but not in sample by the month of interview.

Type Z noninterview

An eligible person in an interviewed household from whom the field representative could not get an interview or for whom the interviewer could not obtain a proxy interview. A noninterview also occurs when a person who was part of the household for a portion of the reference period moves and is no longer a household member on the date of the interview. If the person is an original sample member, an effort will be made to locate and follow the person.

U

undercoverage

Underrepresentation of demographic subgroups within the surveyed population.

unrelated subfamily

A family, that is, a group of two or more related individuals, living at a sample address unit that does not contain the reference person or anyone related to the reference person.

User Notes

Issued periodically by the Census Bureau, these contain updated information for specific microdata files.

usual place of residence

Place where a person normally lives and sleeps; specific living quarters held for the person, to which he or she is free to return at any time.

V

variable metadata

Provides a complete characterization of a variable's content. Variable metadata are available on the SIPP Web site.

W

wave

One round of interviewing, which takes 4 months to complete; one fourth of the sample (i.e., a rotation group) is interviewed each month.

wave files

See *core wave files*.

weights

Estimates of the number of units in the target population that a given unit represents.

X

Y

Z